

A Smart Borewell System for Locating Lost Children via the Use of GPS and AI

V. SHASHIVANTH, N.NAVEEN KUMAR

University College of Engineering, Science and Technology, Hyderabad,
JNTUH.

Associate Professor of IT, Jawaharlal Nehru Technological University,
Hyderabad.

ABSTRACT:

Since the majority of Indian farmers and ranchers depend on agriculture for their daily water requirements, groundwater is essential to the country's economy. If water had given up on its basic need, the drag wells would have been left uncovered. Because of this, most children push towards the well without thinking and fall in. This is a frightening prospect for parents worldwide, since there have been several reports of children being tragically stuck in bore wells. For children to avoid boredom, there should be a great focus on developing a rescue and safety system. Because there is a great chance of human mistake, the current salvage frameworks need manual help, which takes time and is imprecise. The drag well incidents have taken some young, innocent lives while these kids were only having fun. This project intends to include a machine learning algorithm-

powered handling and camera module into the apparatus in order to meet this difficulty. We can identify whether a kid is sad by glancing at their face and using the Accuracy calculation. If the child is, we can play calming music to help them feel better. The Fscore calculation is used to ascertain if the child understands the questions being posed via the audio output. Temperature sensors allow physicians to make an informed assessment as to whether a youngster is still alive. A web application written using ReactJS has been created to better highlight the situation with the child's positive attitude. Consequently, the programme offers a practical, creative approach for the child's protection.

INTRODUCTION

India is a country that relies heavily on horticulture, with most farmers and ranchers

relying on groundwater to meet their basic water needs. Drag wells would have often remained uncovered owing to the necessary precondition yielding of the water. This is the main cause of these ultimately miserable disasters, since the majority of kids unintentionally push towards the well and fall into it. Recently, several reports of children drowning after being trapped in bore wells have been made; for parents, this is a nightmare come true. It is imperative that a security and salvage structure be established for children in order to protect them from bore wells. The salvage frameworks in place now take a long time and are quite inaccurate since they rely on manual help, which increases the possibility of human error. Many innocent children have lost their lives while playing here due to drag well instances. In order to overcome this problem, a camera and handling module for the Raspberry Pi will be used by an equipment module in this assignment. A passionate vggnet computation is used to interpret the child's facial expressions in order to determine whether or not the child is in distress. If the child remains silent, music is played. A sound yield is used to offer questions to the child, and VGGish calculation is used to determine whether or not the child responds appropriately, indicating its level of

cognition. To verify that the child is alive, the temperature of the child is calculated using warm sensors. A reactJS web application is developed to illustrate the child's excited state of affairs. As a result, the project successfully provides a novel solution for child protection.

RELATED WORK

”Context Based Emotion Recognition using EMOTIC Dataset”

In our everyday lives and social interactions we often try to perceive the emotional states of people. There has been a lot of research in providing machines with a similar capacity of recognizing emotions. From a computer vision perspective, most of the previous efforts have been focusing in analyzing the facial expressions and, in some cases, also the body pose. Some of these methods work remarkably well in specific settings. However, their performance is limited in natural, unconstrained environments. Psychological studies show that the scene context, in addition

to facial expression and body pose, provides important information to our perception of people's emotions. However, the processing of the context for automatic emotion recognition has not been explored in depth, partly due to the lack of proper data. In this paper we present EMOTIC, a dataset of images of people in a diverse set of natural situations, annotated with their apparent emotion. The EMOTIC dataset combines two different types of emotion representation: (1) a set of 26 discrete categories, and (2) the continuous dimensions Valence, Arousal, and Dominance. We also present a detailed statistical and algorithmic analysis of the dataset along with annotators' agreement analysis. Using the EMOTIC dataset we train different CNN models for emotion recognition, combining the information of the bounding box containing the person with the contextual information extracted from the scene. Our results show how scene

context provides important information to automatically recognize emotional states and motivate further research in this direction.

“Cross-Subject Emotion Recognition using Flexible Analytic Wavelet Transform from EEG Signals”

Human emotion is a physical or psychological process which is triggered either consciously or unconsciously due to perception of any object or situation. The electroencephalogram (EEG) signals can be used to record ongoing neuronal activities in the brain to get the information about the human emotional state. These complicated neuronal activities in the brain cause non-stationary behavior of the EEG signals. Thus, emotion recognition using EEG signals is a challenging study and it requires advanced signal processing techniques to extract the hidden information of emotions from EEG signals. Due to poor

generalizability of features from EEG signals across subjects, recognizing cross-subject emotion has been difficult. Thus, our aim is to comprehensively investigate the channel specific nature of EEG signals and to provide an effective method based on flexible analytic wavelet transform (FAWT) for recognition of emotion. FAWT decomposes the EEG signal into different sub-band signals. Furthermore, we applied information potential to extract the features from the decomposed sub-band signals of EEG signal. The extracted feature values were smoothed and fed to the random forest and support vector machine classifiers that classified the emotions. The proposed method is applied to two different publicly available databases which are SJTU emotion EEG dataset and database for emotion analysis using physiological signal. The proposed method has shown better performance for human emotion classification as compared to

the existing method. Moreover, it yields channel specific subject classification of emotion EEG signals when exposed to the same stimuli.

“Cross-corpus Acoustic Emotion Recognition with Multi-task Learning: Seeking Common Ground while Preserving Differences”

There is growing interest in emotion recognition due to its potential in many applications. However, a pervasive challenge is the presence of data variability caused by factors such as differences across corpora, speaker's gender, and the “domain” of expression (e.g., whether the expression is spoken or sung). Prior work has addressed this challenge by combining data across corpora and/or genders, or by explicitly controlling for these factors. In this work, we investigate the influence of corpus, domain, and gender on the cross-corpus generalizability of emotion recognition systems. We use a multi-

task learning approach, where we define the tasks according to these factors. We find that incorporating variability caused by corpus, domain, and gender through multi-task learning outperforms approaches that treat the tasks as either identical or independent. Domain is a larger differentiating factor than gender for multi-domain data. When considering only the speech domain, gender and corpus are similarly influential. Defining tasks by gender is more beneficial than by either corpus or corpus and gender for valence, while the opposite holds for activation. On average, cross-corpus performance increases with the number of training corpora. The results demonstrate that effective cross-corpus modeling requires that we understand how emotion expression patterns change as a function of non-emotional factors.

”Heterogeneous Acceleration of HAR Applications”

Human action recognition (HAR) is an important field of research that intercepts with areas such as image processing, computer vision, and the design of fast algorithms, among others.

” Internal Emotion Classification Using EEG Signal with Sparse Discriminative Ensemble”

Among various physiological signal acquisition methods for the study of the human brain, EEG (Electroencephalography) is more effective. EEG provides a convenient, non-intrusive, and accurate way of capturing brain signals in multiple channels at fine temporal resolution. We propose an ensemble learning algorithm for automatically computing the most discriminative subset of EEG channels for internal emotion recognition. Our method describes an EEG channel using kernel-based representations computed from the training EEG recordings. For ensemble learning, we formulate a graph

embedding linear discriminant objective function using the kernel representations. The objective function is efficiently solved via sparse non-negative principal component analysis and the final classifier is learned using the sparse projection coefficients. Our algorithm is useful in reducing the amount of data while improving computational efficiency and classification accuracy at the same time. The experiments on publicly available EEG dataset demonstrate the superiority of the proposed algorithm over the compared methods.

”ISLA: Temporal Segmentation and Labeling for Audio-Visual Emotion Recognition”

Emotion is an essential part of human interaction. Automatic emotion recognition can greatly benefit human-centered interactive technology, since extracted emotion can be used to understand and respond to user needs. However, real-world emotion recognition faces a central challenge

when a user is speaking: facial movements due to speech are often confused with facial movements related to emotion. Recent studies have found that the use of phonetic information can reduce speech-related variability in the lower face region. However, methods to differentiate upper face movements due to emotion and due to speech have been underexplored. This gap leads us to the proposal of the Informed Segmentation and Labeling Approach (ISLA). ISLA uses speech signals that alter the dynamics of the lower and upper face regions. We demonstrate how pitch can be used to improve estimates of emotion from the upper face, and how this estimate can be combined with emotion estimates from the lower face and speech in a multimodal classification system. Our emotion classification results on the IEMOCAP and SAVEE datasets show that ISLA improves overall classification performance. We also

demonstrate how emotion estimates from different modalities correlate with each other, providing insights into the differences between posed and spontaneous expressions.

”Measurement and Analysis of Local Pulse Transit Time for Emotion Recognition”

Emotion recognition based on physiological parameters is a research field that pushes forward from lab settings to real-life investigations. Wearable devices facilitate this advance. However, these devices are still functionally limited compared to stationary medical devices. Our goal is to extend the capability of wearable devices by developing a method that measures the pulse transit time (PTT) locally. The PTT is an interesting parameter regarding emotion recognition. A method for local PTT measurement can be implemented using two photoplethysmography (PPG) sensors. However, this method is error-prone. In this paper, the

physiological background that is presumably responsible for erroneous PPG-based PTT measurements is discussed. We present an algorithm that is capable to handle the derived physiological effects. The algorithm analyzes and compares the two PPG-signals to adapt to time-varying physiological effects. By using this algorithm, calculating and analyzing of the local PTT in the context of emotion recognition become possible. A study (n = 40) to test the algorithm and investigate the usefulness of local PTT analysis for emotion recognition in combination with other physiological signals was conducted. PTT-based parameters, which were derived from the frequency domain of the signal, showed a statistically significant ($p < 0.05$) difference between induced emotional states, if calculated by the developed algorithm. Our findings indicate that parameters derived by our method are significantly affected by emotional stimuli. We suggest that this

method can be used to advance emotion recognition investigations in real life as it can potentially be integrated into a single wearable device.

”Reliable Crowd sourcing and Deep Locality- Preserving Learning for Unconstrained Facial Expression Recognition”

Facial expression is central to human experience, but most previous databases and studies are limited to posed facial behavior under controlled conditions. In this paper, we present a novel facial expression database, Real-world Affective Face Database (RAF-DB), which contains approximately 30 000 facial images with uncontrolled poses and illumination from thousands of individuals of diverse ages and races. During the crowd sourcing annotation, each image is independently labelled by approximately 40 annotators. An expectation-maximization algorithm is developed to reliably estimate the

emotion labels, which reveals that real-world faces often express compound or even mixture emotions. A cross-database study between RAF-DB and CK+ database further indicates that the action units of real-world emotions are much more diverse than, or even deviate from, those of laboratory-controlled emotions. To address the recognition of multi-modal expressions in the wild, we propose a new deep locality-preserving convolutional neural network (DLP-CNN) method that aims to enhance the discriminative power of deep features by preserving the locality closeness while maximizing the inter-class scatter. Benchmark experiments on 7-class basic expressions and 11-class compound expressions, as well as additional experiments on CK+, MMI, and SFEW 2.0 databases, show that the proposed DLP-CNN outperforms the state-of-the-art handcrafted features and deep learning-based methods for expression recognition in the wild. To

promote further study, we have made the RAF database, benchmarks, and descriptor encodings publicly available to the research community.

“Emotion recognition in context,”

Understanding what a person is experiencing from her frame of reference is essential in our everyday life. For this reason, one can think that machines with this type of ability would interact better with people. However, there are no current systems capable of understanding in detail peoples emotional states. Previous research on computer vision to recognize emotions has mainly focused on analyzing the facial expression, usually classifying it into the 6 basic emotions [11]. However, the context plays an important role in emotion perception, and when the context is incorporated, we can infer more emotional states. In this paper we present the Emotions in Context Database (EMCO), a dataset of images containing people in context in non-

controlled environments. In these images, people are annotated with 26 emotional categories and also with the continuous dimensions valence, arousal, and dominance [21]. With the EMCO dataset, we trained a Convolutional Neural Network model that jointly analyzes the person and the whole scene to recognize rich information about emotional states. With this, we show the importance of considering the context for recognizing peoples emotions in images, and provide a benchmark in the task of emotion recognition in visual context.

“Chalearn looking at people: Events and resources,”

This paper reviews the historic of ChaLearn Looking at People (LAP) events. We started in 2011 (with the release of the first Kinect device) to run challenges related to human action/activity and gesture recognition. Since then we have regularly organized events in a series of

competitions covering all aspects of visual analysis of humans. So far we have organized more than 10 international challenges and events in this field. This paper reviews associated events, and introduces the ChaLearn LAP platform where public resources (including code, data and preprints of papers) related to the organized events are available. We also provide a discussion on perspectives of ChaLearn LAP activities.

METHODOLOGY

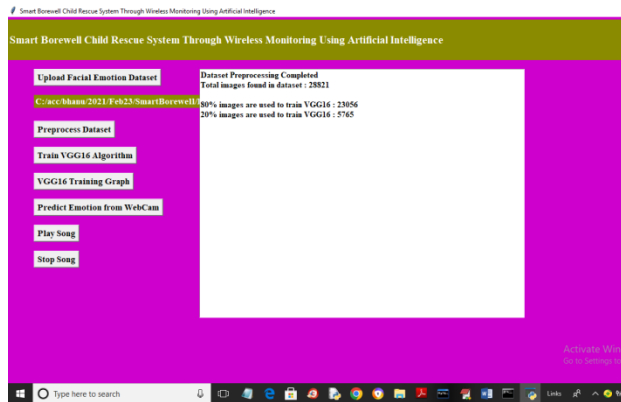
- 1) **Upload Facial Emotion Dataset:** using this module we can upload dataset to application and then find different species found in dataset
- 2) **Preprocess Dataset:** using this module we will read each image and then resize images to equal sizes and then normalized pixel values and then split dataset into train and test where application

used 80% images for training and 20% for testing

- 3) **Train VGG16 Algorithm:** using this module we will input 80% images to train VGG19 model and then apply 20% test images on trained model to calculate prediction accuracy.
- 4) **VGG16 Training Graph:** using this module we will plot VGG16 training accuracy and loss graph
- 5) **Predict Emotion from Webcam:** using this module WEBCAM will be open and then detect faces and then VGG16 will analyse face to predict emotion.
- 6) **Play Song:** based on predicted emotion song will be played

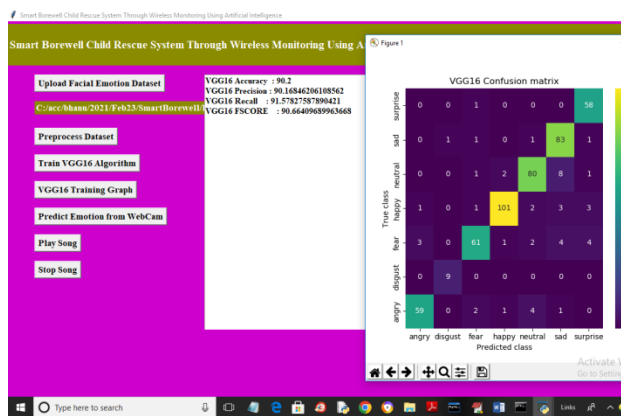
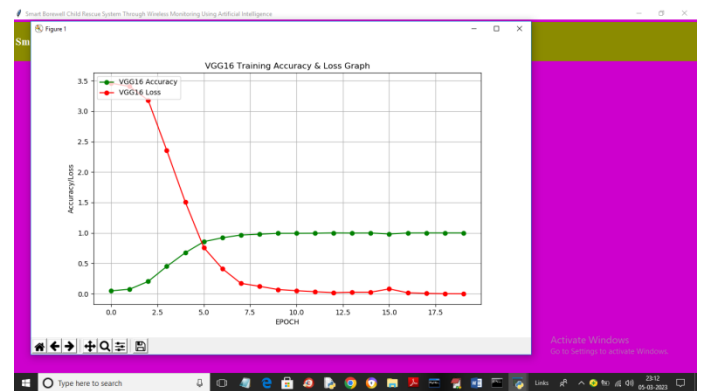
Stop Song: can be used to stop song

RESULT AND DISCUSSION



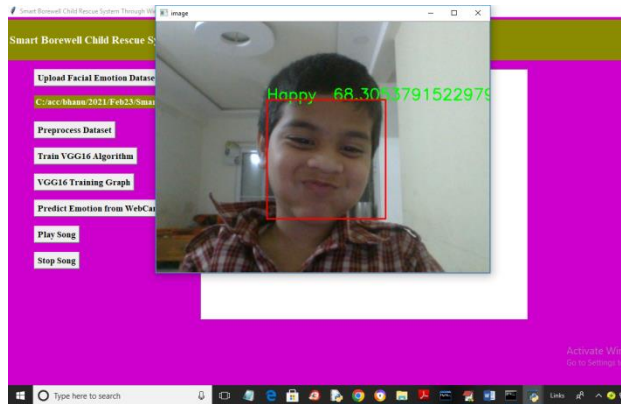
In above screen all images are processed and total images found in dataset is 28821 and application using 23056 (80%) images for training and 5765 (20%) images for testing and now click on 'Train VGG16 Algorithm' button to train VGG16 model and get below output

represents True classes and different colour boxes in Diagnol represents correct prediction count and all other blue boxes contains incorrect prediction count which are very few. Now close above graph and then click on 'VGG16 Training Graph' button to get below VGG16 training graph.



In above screen with VGG16 we got accuracy and precision as 90% and in confusion matrix graph x-axis represents Predicted Labels and y-axis

In above graph x-axis represents training epoch and y-axis represents accuracy and red line represents LOSS values and green line represents accuracy and with each increasing epoch got increased and loss got decreased. Now close above graph and then click on 'Predict Emotion from Webcam' button to upload open Webcam and get below output



In above screen Webcam started and we can see detected emotion with accuracy % and you can press 'q' key to stop webcam and then click on 'Play Song' button to play the song and after some time you can hear song voice and while playing you can click on 'Stop Song' to stop playing. For playing song application will open new window you just closed that newly open window to hear song.

CONCLUSION

This project has been successfully implemented for effectively monitoring and rescuing a child on being stuck in a borewell, the system can effectively recognize the facial emotions and recognize the voice of a

child using the currently prevailing deep learning approach.

REFERENCES

- [1] Ronak Kosti, Jose M. Alvarez, Adria Recasens, Agata Lapedriza."Context Based Emotion Recognition using EMOTIC Dataset", [2019, Vol No: 0162-8828]
- [2] Vipin Gupta, Mayur DahyabhaiChopda, and Ram Bilas Pachori, Senior Member, IEEE, "Cross-Subject Emotion Recognition using Flexible Analytic Wavelet Transform from EEG Signals" [IEEE 2018, Vol .No.: 1558-1748].
- [3] Biqiao Zhang, Student Member, IEEE, Emily Mower Provost, Member, IEEE, and Georg Essl, Member, IEEE, "Cross-corpus Acoustic Emotion Recognition with Multi-task Learning: Seeking Common Ground while Preserving Differences" [2016, Vol.No: 1949-3045].

[4] Jose Rodriguez Borbon, Student Member, IEEE, Xiaoyin Ma, Member, IEEE, Amit K. Roy-Chowdhury, Senior Member, IEEE, and Walid Najjar, Fellow, IEEE, "Heterogeneous Acceleration of HAR Applications"[2018, Vol No: 1051-8215].

[5] Habib Ullah¹, Muhammad Uzair², Arif Mahmood³, Mohib Ullah⁴, Sultan Daud Khan⁵, Faouzi Alaya Cheikh, "Internal Emotion Classification Using EEG Signal with Sparse Discriminative Ensemble"[2018, Vol. No: 2169-3536]

[6] Yelin Kim, Member, IEEE, and Emily Mower Provost, Member, IEEE, "ISLA: Temporal Segmentation and Labeling for Audio-Visual Emotion Recognition"[2017, Vol No: 1949-3045].

[7] N. Beckmann, R. Viga, A. Dogangün, and A. Grabmaier, "Measurement and

Analysis of Local Pulse Transit Time for Emotion Recognition"[2019, Vol. 19, No. 17].

[8] Yong-Jin Liu, Senior Member, IEEE, Minjing Yu, Guozhen Zhao, Jinjing Song, Yan Ge, and Yuanchun Shi, Senior Member, IEEE, "RealTime Movie-Induced Discrete Emotion Recognition from EEG Signals"[2017, Vol No: 1949-3045].

[9] Shan Li, and Weihong Deng, Member, IEEE, "Reliable Crowdsourcing and Deep Locality-Preserving Learning for Unconstrained Facial Expression Recognition"[2018, Vol No: 1057- 7149].

[10] R. Girshick, "Fast r-cnn," in Proceedings of the IEEE international conference on computer vision, 2015, pp. 1440–1448.

[11] R. Kosti, J. M. Alvarez, A. Recasens, and A. Lapedriza, "Emotion recognition in context," in The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.

- [12] Baro, H. J. Escalante, and I. Guyon, ' "Chalearn looking at people: Events and resources," CoRR, vol. abs/1701.02664, 2017.
- [13] T. Lin, M. Maire, S. J. Belongie, L. D. Bourdev, R. B. Girshick, J. Hays, P. Perona, D. Ramanan, P. Dollar, 'and C. L. Zitnick, "Microsoft COCO: common objects in context," CoRR, vol. abs/1405.0312, 2014.
- [14] Toward User-Independent Emotion Recognition using Physiological Signals, Amani Albraikan, Diana P. Tob'on, and Abdulmotaleb El Saddik [2018, Vol. No: 1558-1748]
- [15] W. Mou, O. Celiktutan, and H. Gunes, "Group-level arousal and valence recognition in static images: Face, body and context," in Automatic Face and Gesture Recognition (FG), 2015 11th IEEE International Conference and Workshops on, vol. 5. IEEE, 2015, pp. 1– 6.
- [16] M. B. Reshma A.B., N. Rejee W.H., S. Revathy and J. Jabez, "Recommendation of Nearest Heart care Center during Emergency using IoT," 2019 International Conference on Intelligent Computing and Control Systems (ICCS), Madurai, India, 2019, pp. 1056-1060, doi: 10.1109/ICCS45141.2019.9065409.
- [17] M. B. Reshma A.B., N. Rejee W.H., S. Revathy and J. Jabez, "Recommendation of Nearest Heart care Center during Emergency using IoT," 2019 International Conference on Intelligent Computing and Control Systems (ICCS), Madurai, India, 2019, pp. 1056-1060, doi: 10.1109/ICCS45141.2019.9065409.
- [18] Subhashini, R., and R. Sethuraman. "IoT-Based Air Pollution Monitoring Using Silver Birch Trees." National Academy Science Letters (2019): 1-4.